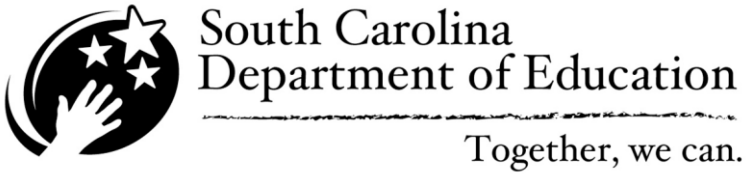
SOUTH CAROLINA
END-OF-COURSE EXAMINATION PROGRAM

2007–08 OPERATIONAL TEST TECHNICAL REPORT



Issued by the

South Carolina Department of Education

Office of Assessment
Division of Accountability

Jim Rex
State Superintendent of Education

CONTENTS

LIST OF TABLES

CHAPTER 1

HISTORY AND OVERVIEW

The South Carolina Education Accountability Act of 1998 requires the development of end-of-course examinations in gateway, or benchmark, courses for grades nine through twelve. When the program is fully implemented, all students enrolled in End-of-Course Examination Program (EOCEP) courses will take the tests for those courses: Algebra 1, Mathematics for the Technologies 2, English 1, Physical Science, Biology 1, Applied Biology 2, and U.S. History and the Constitution.

As they are enunciated in State Board of Education Regulation 43-262.4, the purposes and uses of the EOCEP tests are as follows:

A. The tests shall promote instruction in the specific academic standards for the courses, encourage student achievement, and document the level of students' mastery of the curriculum standards.

B. The tests shall serve as indicators of program, school, and school district effectiveness in the manner prescribed by the Education Oversight Committee in accordance with the provisions of the Education Accountability Act of 1998 (EAA).

C. The tests shall be weighted 20 percent in the determination of students' final grades in the gateway courses.

EOCEP exams will be reported on the basis of the South Carolina uniform grading scale (UGS). The score reported is a scale score and not the percentage of correct answers.

The Algebra 1/Mathematics for the Technologies 2 end-of-course examination was implemented in the baseline year 2002–03 and was operational for the first time in 2003–04. The English 1, Physical Science, and Biology 1/Applied Biology 2 examinations that were field-tested in May 2003 were implemented for the baseline year in 2003–04. These subject-area EOCEP examinations became operational in 2004–05. The Biology 1/Applied Biology 2 examination was discontinued after the 2005–06 school year. However, the State Board of Education decided to reinstate the Biology test, beginning in the spring of 2008 with a field test. The U.S. History and the Constitution examination was field-tested in 2005–06, with baseline implementation in 2006–07, and it was scheduled to be operationally administered in 2007–08. However, it was determined that the U.S. History and the Constitution examination needed an additional year of implementation and, thus, was not operational in 2007-08.

The South Carolina Department of Education (SCDE) awarded the contract for the development and scoring of the EOCEP tests in October 2001 to American Institutes for Research (AIR) and its partners Insite, Inc., and Pearson Educational Measurement (PEM). These contractors have undertaken a number of development, review, implementation, and data analysis activities. In spring 2007, Pearson became the sole contractor.

All EOCEP exams contain only multiple-choice items. For all subjects, with the exception of U.S. History and the Constitution, the tests were operational in 2007–08. Rasch-ability-score-to-

scale-score conversion tables were produced prior to each test administration on the basis of the item parameters in the pre-equated item pool. This technical report summarizes the results of statistical and psychometric analyses performed on the current year's operational data.

In this report, all data are based on the students in the regular schools and in adult education programs only. Data on students in district-approved homeschools have been excluded.

CHAPTER 2

STUDENT DEMOGRAPHICS

2.1 STUDENT PARTICIPATION

All schools administered EOCEP tests to the students who completed courses for Algebra 1, Mathematics for the Technologies 2, Biology 1, Applied Biology 2, Physical Science, U.S. History and the Constitution, or English 1 for credit toward a high school diploma. Summary data are reported for operational tests only.

Demographic data were collected for each student. These data included the categories of gender, race/ethnicity, grade, English language fluency (LEP, limited English proficiency), lunch program participation, individualized education program (IEP) status, disability status, and migrant status. Table 2.1 presents the combined student participation in the three EOCEP administrations (fall, spring, and summer) by the demographic variables.

Table 2.1

Summary of Student Demographics in the Sample

Demographics	Algebra 1/ Math Tech 2		English 1		Physical Science	
	N	%	N	%	N	%
Overall	60,125		58,433		50,259	
Gender						
Female	29,969	49.84	28,729	49.17	24,945	49.63
Male	29,902	49.73	29,466	50.43	25,183	50.11
Unknown	254	0.42	238	0.41	131	0.26
Grade						
6	9	0.01	0	0.00	0	0.00
7	2,434	4.05	5	0.01	0	0.00
8	12,879	21.42	8,274	14.16	418	0.83
9	30,142	50.13	48,688	83.32	36,932	73.48
10	12,805	21.30	885	1.51	11,431	22.74
11	1,199	1.99	136	0.23	855	1.70
12	268	0.45	24	0.04	313	0.62
Adult education	44	0.07	14	0.02	30	0.06
Other	345	0.57	407	0.70	280	0.56
Ethnicity						
White	32,638	54.28	31,412	53.76	27,766	55.25
African-American	22,967	38.20	22,668	38.79	18,865	37.54
Hispanic	2,430	4.04	2,452	4.20	2,031	4.04
Asian	656	1.09	597	1.02	518	1.03
Hawaiin/Pacific Islander	65	0.11	46	0.08	50	0.10
American Indian	102	0.17	120	0.21	115	0.23
Other	1,010	1.68	888	1.52	779	1.55
Unknown	257	0.43	250	0.43	135	0.27

Table 2.1

Summary of Student Demographics in the Sample

Language						
Parent waiver	93	0.15	72	0.12	75	0.15
Pre-functional	265	0.44	288	0.49	216	0.43
Beginner	222	0.37	247	0.42	207	0.41
Intermediate	327	0.54	324	0.55	257	0.51
Advanced	511	0.85	476	0.81	352	0.70
Initially English Proficient	173	0.29	134	0.23	104	0.21
Title III First Year Exited	2	0.00	2	0.00	1	0.00
Title III Second + Year Exited	222	0.37	212	0.36	203	0.40
English Speaker I	239	0.40	225	0.39	171	0.34
English Speaker II	55,367	92.09	54,239	92.82	47,038	93.59
Unknown	2,704	4.50	2,214	3.79	1,635	3.25
Lunch						
Free meals	22,282	37.06	22,801	39.02	18,423	36.66
Reduced-price meals	4,405	7.33	4,392	7.52	3,600	7.16
No free/reduced-price meals	31,960	53.16	30,037	51.40	27,026	53.77
Unknown	1,478	2.46	1,203	2.06	1,210	2.41
IEP						
Yes	4,641	7.72	5,032	8.61	4,240	8.44
No	55,484	92.28	53,401	91.39	46,019	91.56
Migrant						
Yes	28	0.05	41	0.07	28	0.06
No	6,714	11.17	6,292	10.77	5,404	10.75
Unknown	53,383	88.78	52,100	89.16	44,827	89.19
Gifted/talented						
Academic	48,087	79.98	47,778	81.77	43,188	85.93
Artistic	10,120	16.83	9,054	15.49	5,976	11.89
Both	346	0.58	351	0.60	343	0.68
No	1,208	2.01	990	1.69	622	1.24
Unknown	364	0.61	260	0.44	130	0.26
504 plan						
Yes	642	1.07	629	1.08	529	1.05
No	40,671	67.64	40,424	69.18	35,605	70.84
Unknown	18,812	31.29	17,380	29.74	14,125	28.10
Alternative school						
Yes	947	1.58	1,328	2.27	932	1.85
No	17,869	29.72	17,078	29.23	15,143	30.13
Unknown	41,309	68.71	40,027	68.50	34,184	68.02
Accommodations						
Yes	2,281	3.79	2,508	4.29	2,235	4.45
No	57,844	96.21	55,925	95.71	48,024	95.55

2.2 ACCOMMODATIONS

Supplemental information regarding the administration of the EOCEP to students with disabilities is contained in the EOCEP test administration manuals (SCDE 2007b, 2008b, and

2008d). These manuals provide guidelines for IEP teams in making decisions about testing students with disabilities and gives specific information regarding testing accommodations and modifications, test forms and materials, and test administration procedures.

A student with a documented disability is one who has been evaluated and found to meet the eligibility criteria for enrollment in special education as defined by the 1997 amendments to the Individuals with Disabilities Education Act and by State Board of Education Regulation 43-243.1, or one who has a disability covered under Section 504 of the Rehabilitation Act of 1973. The IEP or 504 plan team determines how a student with disabilities participates in the EOCEP assessments. Decisions about accommodations and modifications must be made on an individual student basis, not on the basis of the category of disability. Table 2.2 presents the percentages of accommodations used in the current year's testing.

Table 2.2
Accommodations Used in 2007–08 EOCEP Testing

Accommodations	Algebra 1/ Math Tech 2	English 1	Physical Science
	Regular Form		
	(N = 58,394)	**(N = 56,391)**	**(N = 48,463)**
Setting	1.45	1.57	1.49
Timing	4.86	6.02	5.01
Scheduling	2.2	1.9	4.6
Response options	1.39	1.11	1.22
Presentation	1.62	1.9	2.03
	Customized Form		
	(N = 1,731)	**(N = 2,042)**	**(N = 1,796)**
Setting	73.19	70.52	74.72
Timing	4.73	7.51	6.15
Scheduling	2.61	5.28	3.48
Response options	1.91	2.79	1.6
Presentation	62.31	55.62	61.56

Total responses in each column may exceed 100 percent because some students received accommodations in more than one category.

2.3 TEST ADMINISTRATION TIME

In addition to providing their demographic information, students were asked to record on their answer documents the exact times that they started and finished the test. These answer documents were scanned, and the total elapsed time was calculated for each student. (It was not possible to calculate a total testing time for students with incomplete or invalid data.) A large majority of students finished the test within two hours, as tables 2.3 and 2.4 reflect.

Table 2.3
Time Taken in 2007–08 EOCEP Testing with Regular Forms

	Algebra 1/Math Tech 2			English 1		
	Fall 2007 (N = 8,297)	Spring 2008 (N = 48,646)	Summer 2008 (N = 612)	Fall 2007 (N = 7,593)	Spring 2008 (N = 46,388)	Summer 2008 (N = 394)
30 min	0.43	0.37	1.29	0	0.33	0.73
45 min	2.12	2.07	5.82	1.26	1.11	3.65
1 hr	9.32	10.15	20.19	5.99	6.59	13.14
1 hr 15 min	19.56	21.71	22.94	15.52	18.3	20.68
1 hr 30 min	24.46	23.36	19.06	22.97	22.33	18.73
1 hr 45 min	17.99	17.09	11.95	19.21	18.43	14.6
2 hr	10.5	9.77	6.3	12.3	12.16	9.98
2 hr 15 min	5.07	5.26	4.2	8.5	7.68	4.87
2 hr 30 min	2.61	2.6	2.75	4.35	0	0
2 hr 45 min	1.19	1.48	0.97	2.28	2.07	0.73
3 hr or more	0.63	0.8	0	1.01	1.14	2.92
Invalid*	4.89	3.86	3.39	6.09	5.91	6.57

	Physical Science		
	Fall 2007 (N = 12,605)	Spring 2008 (N = 32,436)	Summer 2008 (N = 158)
30 min	0.93	1.23	11.39
45 min	13.16	14.88	32.28
1 hr	33.95	35.62	24.68
1 hr 15 min	29.84	27.41	17.09
1 hr 30 min	11.96	10.13	5.7
1 hr 45 min	3.89	3.3	1.27
2 hr	1.2	1.47	1.27
2 hr 15 min	0.6	0.67	0
2 hr 30 min	0	0.28	6.33
2 hr 45 min	0.09	0.04	
3 hr or more	0.22	0.19	
Invalid*	3.94	4.66	

* includes responses with no mark or multiple marks on start and/or stop time fields, making it impossible to compute the difference between start and stop times

Table 2.4

Time Taken in 2007–08 EOCEP Testing with Regular Forms

	Algebra 1/Math Tech 2			English 1		
	Fall 2007 (N = 368)	Spring 2008 (N = 1,324)	Summer 2008 (N = 4)	Fall 2007 (N =307)	Spring 2008 (N = 1,7138)	Summer 2008 (N = 12)
15 min		0.52	50		0.17	
30 mn	2.17		50	4.22		15.38
45 min	6.25	7.73		11.04	3.2	30.77
1 hr	16.85	19.87		17.53	7.21	38.46
1 hr 15 min	15.49	17.37		23.05	16.21	7.69
1 hr 30 min	16.85	17.37		13.64	24.35	
1 hr 45 min	10.05	10.38		11.04	16.27	
2 hr	10.33	6.25		3.25	10.98	
2 hr 15 min	4.89	3.75		1.62	5.81	
2 hr 30 min	3.53	2.5		1.62	1.92	
2 hr 45 min	1.36	0.59		12.66	0.81	
3 hr or more	2.17	1.62			1.63	
Invalid*	10.05	9.49			10.98	

	Physical Science		
	Fall 2007 (N = 409)	Spring 2008 (N = 1,343)	Summer 2008 (N = 1)
15 min	0	0.51	100
30 mn	1.96		
45 min	9.29	13.06	
1 hr	29.83	21.28	
1 hr 15 min	19.07	23.3	
1 hr 30 min	13.2	13.92	
1 hr 45 min	8.8	7.65	
2 hr	4.16	3.46	
2 hr 15 min	0.98	1.3	
2 hr 30 min	0.98	0.22	
2 hr 45 min	0.73	0.72	
3 hr or more	1.96	1.15	
Invalid*	9.05	10.32	

* includes responses with no mark or multiple marks on start and/or stop time fields, making it impossible to compute the difference between start and stop times

2.4 STUDENT QUESTIONNAIRE

After the administration of the EOCEP test in each subject, students were instructed to complete a questionnaire that addressed such topics as the difficulty of the test, the nature of the instruction they had received in the particular course, their use of calculators in the particular course (algebra only), and the amount of time they had spent engaged in lab activities in the particular course (biology and physical science only).

CHAPTER 3

TEST ADMINISTRATION

3.1 TEST ADMINISTRATION WINDOW

The test administration dates for the current year are given in table 3, below. School districts were required to administer all EOCEP tests within a single five-day period. Districts were instructed to administer makeup tests following their regular testing period. For all three EOCEP administrations, district test coordinators (DTCs) were responsible for providing the testing schedule to all school test coordinators (STCs) in their particular districts.

For students who missed the originally scheduled EOCEP test due to a death in the family, illness, or another situation deemed valid by the state, school districts were required to have a five-day makeup period the week immediately following the original test administration. It was recommended that a single makeup test be given per day, but two could have been given per day if necessary.

TABLE 3

2007–08 EOCEP Test Administration Windows

Administration	Dates
Fall 2007	December 10, 2007–January 25, 2008
Spring 2008	May 5, 2008–June 6, 2008
Summer 2008	June 23, 2008–July 31, 2008

3.2 TIMING OF THE TEST

The EOCEP tests were not timed; however, each session had to be administered during a single day (unless a student's IEP or 504 plan specifically stated that he or she needed to have the test administered over several days). To ensure an accurate assessment, districts and schools were instructed that students should be given as much uninterrupted time as they needed to complete the test.

3.3 ADMINISTRATION MANUALS

Working with the SCDE, AIR staff drafted the administration manuals for the test. SCDE staff reviewed and revised the manuals, and the AIR finalized and printed them. The EOCEP district test coordinator supplements (SCDE 2007a, 2008a, and 2008c) and the EOCEP test administration manuals (TAMs) were produced for each administration of the EOCEP. The DTC supplements included only the information that DTCs needed for the administration of the EOCEP tests. The TAMs contained the information that STCs, test administrators (TAs), and monitors needed to administer the tests to students in their schools.

The TAMs and the supplements included logistical and administration procedures as well as the directions (scripts) for administering the tests. The DTCs, STCs, and TAs were encouraged to use a form provided in the manuals to offer comments and suggestions on the procedures therein. The comments were compiled in a spreadsheet and sent to the SCDE to review and to use as the basis for potential changes in test procedures. The TAMs also included a testing irregularity form that test administrators were instructed to use to report any problems or deviations from established testing procedures.

Appendix C in the TAMs includes a detailed description of materials available, as well as additional graphics for completing student demographic information and returning scorable and nonscorable test materials. Tables showing the types of customized materials available for students who require such special testing formats were also provided.

3.4 CUSTOMIZED MATERIALS

Customized formats of the EOCEP test were available for Algebra 1/Mathematics for the Technologies 2, Biology 1/Applied Biology 2, Physical Science, and English 1:

- Loose-leaf test booklets—printed single-sided, one item to a page, and bound in three-ring binders—allowed individuals to remove the pages, if necessary, during testing.
- Large-print booklets were produced for students who have difficulty reading text in a standard-size font. The large-print version used an 18-point sans serif font and was issued as a 9 x 12-inch spiral-bound booklet.
- Braille booklets were produced for students who typically read classroom materials in braille. The braille version was issued as spiral-bound booklet containing 11½ x 11-inch interpoint braille pages.
- A regular print Form C test booklet was provided in test packets for students or TAs to use with customized formats such as the oral script, braille, large-print, loose-leaf, and sign language versions. These booklets were saddle-stitched and printed in a 12-point font, just as the regular, noncustomized test booklets were.
- For students whose IEP or 504 plan requires the oral administration of tests, oral administration scripts gave specific directions to TAs regarding the appropriate way to read the test questions, the passages on which the questions were based, and the answer choices.

  Beginning in spring 2005, audiocassettes were also produced to be used in the oral administration of the tests. These audiocassettes contained the directions for administering the tests, the passages that were the basis of the questions, the test questions, and the answer choices. The audiocassettes and the oral administration scripts contained the same information.
- Sign language videotapes—produced for Algebra 1/Mathematics for the Technologies 2, English 1, Biology 1/Applied Biology 2, and Physical Science—included the signed test directions, questions, and response options. The videotapes were produced in two languages: American Sign Language and Pidgin Signed English.

3.5 MATERIALS SHIPPING AND RETURN

For all three administrations, test materials were shipped to district offices approximately two weeks before testing—in time for the DTCs to be able to distribute school materials at least one week before the schools' test dates. Each school's shipment was boxed individually and labeled with the total number of boxes shipped to that school.

The district office was also sent a shipment of noncustomized overage materials, which were to be used by the DTCs to complete any additional materials requests from the STCs. Materials in customized formats were sent only to the schools and only in the quantities ordered.

TAs were instructed to return their test materials to the STCs immediately after the test administration. The STCs then redistributed test materials to the TAs who needed them in order to administer makeup tests. Those TAs were instructed to return the makeup test materials to their STC immediately after the makeup session. DTCs were to arrange for the pickup of all scorable materials for return to PEARSON within three days after testing.

Because the test scores were required to be reported back to the schools quickly for calculating final course grades, a rapid scoring and reporting process was utilized for all three administrations. Each school district could return the scorable materials to PEARSON, in as many as five separate shipments, as they arrived from the schools. Nonscorable materials were to be returned in one shipment within three days of the completion of makeup tests. For all three administrations, step-by-step instructions for returning scorable and nonscorable materials were included in the district materials. These instructions listed the toll-free phone numbers of the trucking companies that the DTCs were instructed to call to schedule pickups of return materials

3.6 TEST SECURITY

Test security is an important issue before, during, and following test administrations. The specific procedures to be followed during the EOCEP test administrations are outlined in the *Test Administration Manual* (SCDE 2007b). Reprinted in the manual are an excerpt from Section 59-1-445 of the South Carolina Code of Laws, a summary of Section 59-1-447 of the Code of Laws, and the entirety of State Board of Education Regulation 43-100.

Section 59-1-445 states in part:

> It is unlawful for anyone knowingly and wilfully [*sic*] to violate security procedures regulations promulgated by the State Board of Education for mandatory tests administered by or through the State Board of Education to students or educators, or knowingly and willfully to:
> (a) Give examinees access to test questions prior to testing;
> (b) Copy, reproduce, or use in any manner inconsistent with test security regulations all or any portion of any secure test booklet;
> (c) Coach examinees during testing or alter or interfere with examinees' responses in any way;
> (d) Make answer keys available to examinees;

(e) Fail to follow security regulations for distribution and return of secure test [materials] as directed, or fail to account for all secure test materials before, during, and after testing;
(f) Participate in, direct, aid, counsel, assist in, encourage, or fail to report any of the acts prohibited in this section.

Regulation 43-100 mandates that "Each local school board must develop and adopt a district test security policy" with procedures for the storage and handling of all test materials and that each district superintendent must annually designate a DTC. The regulation and the *TAM* provide specific security guidelines regarding various aspects of the test administration process (e.g., the storage and handling of test materials, the responsibility of administrators to monitor students during testing and to remove supplemental materials from the testing room, and the requirement that administrators refrain from interference with student responses).

Following the test administration and the return of materials, the DRC generated a missing-document report, listing the identification numbers of unreturned secure materials. The report was used to notify districts of missing materials. A toll-free telephone line was manned to answer questions regarding missing documents, and follow-up procedures were employed until all materials were accounted for. Subsequently, the districts located and returned the materials or sent signed statements indicating that all secure materials had been returned.

Secure Materials

Secure materials—each assigned a human- and machine-readable security identification number—are test booklets, answer documents, customized test materials, and secure administration manuals. Secure materials were locked in storage until the day of the test administration and were signed out when they were to be used, and signed in when they were returned. These materials were not to be left unattended at any time.

CHAPTER 4

TECHNICAL CHARACTERISTICS OF ITEMS

This chapter reports the results of item analyses based on classical test theory (CTT) using a proprietary program designed by the AIR. Item difficulty (*p*) is the proportion (or percentage) of examinees correctly answering a dichotomously scored item.

Item discrimination is defined as a correlation between the item score and the total score. For the discrimination index, point-biserial correlations were produced. In computing the point-biserial correlation, the AIR corrected for spuriousness. In the recoding of missing data for item analysis, all omitted and not-reached items were recoded as incorrect, with a zero score. After discussions between the SCDE and the AIR, it was decided to exclude from the CTT item analyses and item calibrations those students who had used customized test materials.

4.1 ITEM NONRESPONSE RATES

Although the EOCEP tests were not timed, students were required to finish each test during one school day, unless they had an IEP that allowed for accommodations in administration. Districts and schools were instructed that, if they had space and staff available, students should be given as much uninterrupted time as necessary to take the test to ensure an accurate assessment.

The item nonresponse rates indicate the percentage of students who did not reach a particular item and all items thereafter. The item omit rates indicate the percentage of students who did not respond to that particular item but did respond to a later item. The percentages for not-reached and omit rates were quite low—less than 1 percent—in all subjects. These data indicate that students were given ample time to complete the test in every subject.

4.2 CLASSICAL ITEM STATISTICS

Table 4, on the following page, provides a summary of item *p*-values and item discrimination values for operational items for all three administrations.

Table 4
Summary of Classical Item Statistics

Administration	Number of items	*p*-value	Adjusted Point-Biserial Correlation
Algebra 1/Math Tech 2			
Fall 2007	50	0.52	0.32
Spring 2008	50	0.57	0.38
Summer 2008	50	0.49	0.31
English 1			
Fall 2007	55	0.66	0.37
Spring 2008	52*	0.68	0.39
Summer 2008	55	0.59	0.35
Physical Science			
Fall 2007	55	0.56	0.4
Spring 2008	55	0.54	0.35
Summer 2008	55	0.49	0.32

* The spring English form contained 52 items.

CHAPTER 5

ITEM CALIBRATION AND SCALING

5.1 METHODOLOGY AND SOFTWARE

The one-parameter Rasch model (Rasch 1960; Wright and Stone 1979) was used to calibrate all items, using WINSTEPS software (see Linacre and Wright 2003). The WINSTEPS program employs joint maximum likelihood estimation, an approach that estimates the item and person parameters simultaneously.

5.2 ITEM CALIBRATION AND PRE-EQUATING

The AIR conducted field tests with a sufficient number of items to create precalibrated item pools and to construct pre-equated operational-test forms for all tests. For all subjects, the Rasch-ability-score-to-scale-score conversion tables were produced prior to each test administration based on the item parameters in the pre-equated item pools.

5.3 SCALING

The SCDE provided the AIR with initial Rasch-ability-score-to-scale-score conversion tables that showed the transformation of the ability score interval for each scale score for each subject area. The AIR then applied these tables specifically to each test form for each subject area on the basis of the pre-equated item pool. The conversion tables took into account any differences in the difficulty of the various forms. All items shared a common metric so that the scale scores developed for each form were automatically adjusted for differences in item difficulty. For all EOCEP test subjects, the scale scores are now reported according to the South Carolina UGS. Scale scores range from 0 to 100 with a minimum passing score of 70. Each scale score is assigned a letter-grade equivalent (A, B, C, D, or F) in accordance with the UGS.

5.4 DEFINITION OF SCOREABILITY

A student was considered "tested" if the student answered at least one question in the answer document. All tested students' item responses were scored. All omits and not-reached items were recoded as incorrect, with a zero score.

5.5 REPORTING OF ZERO AND PERFECT SCORES

In item response theory (IRT), zero and perfect scores are assigned the ability of minus and plus infinity. The AIR used the WINSTEPS default setting in estimating finite values for the extreme scores. In other words, a fractional score point value was subtracted from perfect scores, and was added to zero scores. The WINSTEPS default value for adjusting the extreme scores for extreme measures is 0.3.

5.6 PERCENTAGE OF STUDENTS SCORING IN EACH LETTER-GRADE EQUIVALENT

Tables 5.1 through 5.8 report student performance for all administrations combined. The results are summarized separately for regular schools and for adult education programs. The number and percentage of students in each letter-grade equivalent and the mean scale score are reported for the test-takers overall and by demographic category.

Table 5.1
Algebra 1/Math Tech 2 Operational Test, Regular Schools:
Percentages of Student Scores in Letter-Grade Equivalents, Overall and by Demographics

Demographics	N	Mean Score	A	B	C	D	F
Overall	60,081	78.99	14.45	16.81	22.86	24.05	21.83
Gender							
Unknown	254	70.63	3.15	7.48	16.93	29.92	42.52
Female	29,948	78.98	14.23	16.33	22.99	24.98	21.47
Male	29879	79.08	14.77	17.36	22.78	23.07	22.02
Grade							
6	9	92	66.67	11.11	11.11	0	11.11
7	2,434	92.93	58.26	23.99	13.48	3.12	1.15
8	12,879	88.09	36.63	27.81	20.81	10.82	3.94
9	30,142	77.03	7.59	15.98	25.06	27.23	24.14
10	12,805	72.74	1.58	7.5	21.65	33.45	35.82
11	1,199	72.32	2.59	7.51	22.77	27.52	39.62
12	268	74.76	3.73	13.81	24.63	26.87	30.97
Other	345	70.21	2.9	6.96	17.68	25.8	46.67
Ethnicity							
White	32,621	82.37	21.02	21.41	24.17	19.9	13.5
African American	22,942	74.15	4.83	10.18	21.14	30.37	33.48
American Indian	102	79.11	13.73	18.63	24.51	24.51	18.63
Asian	656	88.14	42.68	21.19	17.23	12.35	6.55
Hawaiin/Pacific Islander	65	82.55	10.77	35.38	30.77	9.23	13.85
Hispanic	2430	77.56	10.45	16.09	23.62	24.32	25.51
Other	1009	79.18	15.26	18.33	21.8	21.41	23.19
Unknown	256	70.92	3.13	7.81	18.75	28.52	41.8
Language							
Pre-functional	265	69.2	1.51	6.42	10.57	24.91	56.6
Beginner	222	71.65	4.5	6.76	13.51	26.58	48.65
Intermediate	327	75.02	4.59	10.09	25.99	29.36	29.97
Advanced	508	81.87	19.69	19.88	24.61	22.64	13.19

Table 5.1
Algebra 1/Math Tech 2 Operational Test, Regular Schools:
Percentages of Student Scores in Letter-Grade Equivalents, Overall and by Demographics

Initially English Proficient	171	79.22	21.05	14.04	17.54	23.98	23.39
Title III First Year Exited	2	68	0	0	0	50	50
Title III Second + Year Exited	222	80.57	11.26	22.97	29.28	23.42	13.06
English Speaker I	235	81.77	30.21	11.06	19.15	16.6	22.98
English Speaker II	55351	79.35	14.89	17.31	23.17	24.01	20.61
Parent Waiver	92	81.23	15.22	19.57	29.35	22.83	13.04
Unknown	2686	72.63	6.14	8.53	17.61	24.94	42.78
Lunch							
No Meals	31,931	82.39	21.6	21.08	23.71	19.63	13.98
Free Meals	22,274	74.73	5.68	11.28	21.28	29.88	31.87
Reduced Meals	4,405	78.09	10.26	16.57	25.79	25.86	21.52
Unknown	1,471	72.52	4.62	8.43	19.51	26.31	41.13
IEP							
Yes	4,637	70.49	2.37	5.59	15.18	27.52	49.34
No	55,444	79.7	15.46	17.74	23.5	23.76	19.53
Migrant							
No	6701	74.4	7.61	10.88	19.13	26.26	36.11
Yes	28	72.5	0	10.71	25	21.43	42.86
Unknown	53,352	79.57	15.32	17.55	23.33	23.77	20.03
Courses taken							
3142 (Math Tech 2)	14,128	72.6	1.4	7.16	21.39	33.61	36.44
4111 (Alg 1)	43,868	81.09	18.46	20.11	23.52	21.06	16.85
Other	2,085	78.25	18.56	12.66	18.99	22.16	27.63
Gifted/talented							
Academic	10,115	90.86	47.65	27.85	16.52	6.22	1.76
Artistic	346	81.87	17.05	22.54	25.43	23.99	10.98
Both	364	91.51	53.3	21.15	18.13	4.67	2.75
No	48,055	76.54	7.37	14.65	24.31	27.87	25.79
Unknown	1,201	72.52	5.41	6.91	18.9	27.14	41.63
504 plan							
No	40,658	78.98	14.53	16.57	22.91	24.28	21.72
Yes	642	78.28	10.75	16.04	26.32	23.21	23.68
Unknown	18,781	79.04	14.4	17.35	22.65	23.59	22.02
Alternative school							
No	17,857	77.96	12.2	15.27	22.96	25.68	23.89
Yes	946	67.59	1.06	3.28	10.15	24.31	61.21
Unknown	41,278	79.7	15.73	17.78	23.11	23.34	20.04
Accommodations							
No	57,804	79.37	14.97	17.29	23.24	23.95	20.56
Yes	2,277	69.43	1.36	4.39	13.31	26.7	54.24

TABLE 5.2
Algebra 1/Math Tech 2 Operational Test, Adult Education Programs: Percentages of Student Scores in Letter-Grade Equivalents, Overall and by Demographics

Demographics	N	Mean Score	A	B	C	D	F
Overall	44	69.45	0	4.55	20.45	18.18	56.82
Gender							
Female	21	69.48	0	9.52	14.29	19.05	57.14
Male	23	69.43	0	0	26.09	17.39	56.52
Ethnicity							
White	17	73.88	0	11.76	29.41	23.53	35.29
African American	25	66.6	0	0	16	12	72
Other	1	75	0	0	0	100	0
Unknown	1	60	0	0	0	0	100
Language							
Advanced	3	69.33	0	0	33.33	0	66.67
Initially English Proficient	2	64	0	0	0	0	100
English Speaker I	4	76	0	25	25	50	0
English Speaker II	16	71.56	0	6.25	25	18.75	50
Parent Waiver	1	64	0	0	0	0	100
Unknown	18	67.06	0	0	16.67	16.67	66.67
Lunch							
No Meals	29	71.69	0	6.9	24.14	20.69	48.28
Free Meals	8	68.63	0	0	12.5	25	62.5
Unknown	7	61.14	0	0	14.29	0	85.71
IEP							
Yes	4	68.5	0	0	0	25	75
No	40	69.55	0	5	22.5	17.5	55
Migrant							
No	13	71.23	0	0	23.08	23.08	53.85
Unknown	31	68.71	0	6.45	19.35	16.13	58.06
Courses taken							
3142 (Math Tech 2)	17	70.65	0	5.88	23.53	17.65	52.94
4111 (Alg 1)	11	73	0	9.09	27.27	27.27	36.36
Other	16	65.75	0	0	12.5	12.5	75
Gifted/talented							
No	32	71.09	0	6.25	21.88	18.75	53.13
Academic	5	70.6	0	0	20	40	40
Unknown	7	61.14	0	0	14.29	0	85.71
504 plan							
No	13	70.62	0	0	23.08	23.08	53.85
Unknown	31	68.97	0	6.45	19.35	16.13	58.06
Alternative school							
No	12	71.42	0	0	25	25	50
Yes	1	69	0	0	0	0	100
Unknown	31	68.71	0	6.45	19.35	16.13	58.06
Accommodations							
No	40	69.55	0	5	22.5	17.5	55
Yes	4	68.5	0	0	0	25	75

TABLE 5.3
English 1 Operational Test, Regular Schools:
Percentages of Student Scores in Letter-Grade Equivalents, Overall and by Demographics

Demographics	N	Mean Score	A	B	C	D	F
Overall	58,419	76.15	11.86	15.57	23.57	17.26	31.72
Gender							
Unknown	238	66.65	2.94	7.56	15.55	15.13	58.82
Female	28,724	77.41	13.06	16.56	24.7	17.51	28.16
Male	29457	75	10.77	14.67	22.54	17.04	34.98
Grade							
7	5	59.4	0	0	0	0	100
8	8,274	87.15	30.94	30.22	26.78	8.63	3.43
9	48,688	74.59	8.9	13.38	23.29	18.82	35.61
10	885	66.08	2.37	5.65	13.9	15.14	62.94
11	136	65.03	1.47	7.35	12.5	11.03	67.65
12	24	70.88	4.17	16.67	12.5	20.83	45.83
Other	407	64.99	3.44	5.16	17.69	13.02	60.69
Ethnicity							
Unknown	250	66.7	2.8	6.8	18	14	58.4
African American	22,661	70.56	3.25	8.03	20.44	20.82	47.46
American Indian	120	76.12	9.17	14.17	24.17	22.5	30
Asian	597	82.27	24.29	21.27	25.29	11.22	17.92
Hawaiin/Pacific Islander	46	81.41	23.91	15.22	23.91	17.39	19.57
Hispanic	2451	71.59	4.94	11.51	21.66	17.91	43.98
White	31407	80.48	18.46	21.29	25.98	14.74	19.53
Other	887	76.41	11.39	15.9	24.13	18.49	30.1
Language							
Unknown	2,210	66.71	4.71	6.15	13.17	13.44	62.53
Pre-functional	288	56.53	0.35	0	3.13	2.78	93.75
Beginner	247	59.85	0.4	0.4	3.24	9.72	86.23
Intermediate	324	66.59	0.31	2.47	8.33	20.06	68.83
Advanced	476	76.2	6.09	14.29	27.94	24.79	26.89
Initially English Proficient	134	74.95	12.69	14.18	20.9	20.15	32.09
Title III First Year Exited	2	82.5	0	50	50	0	0
Title III Second + Year Exited	212	77.74	10.85	16.04	32.55	16.04	24.53
English Speaker I	224	75.88	13.39	13.39	24.55	17.41	31.25
English Speaker II	54,230	76.77	12.4	16.21	24.22	17.45	29.73

TABLE 5.3
English 1 Operational Test, Regular Schools:
Percentages of Student Scores in Letter-Grade Equivalents, Overall and by Demographics

Parent Waiver	72	71.89	4.17	11.11	26.39	13.89	44.44
Lunch							
Unknown	1,199	67.38	4.25	6.59	14.51	16.68	57.96
No Meals	30,033	80.84	19.2	21.59	25.82	14.56	18.84
Free Meals	22,795	70.67	3.48	8.52	20.5	20.29	47.21
Reduced Meals	4,392	74.93	7.26	13.52	26.66	20.22	32.33
IEP							
Yes	5,031	63.7	1.41	2.98	9.9	14.27	71.44
No	53,388	77.32	12.85	16.76	24.86	17.55	27.98
Migrant							
Unknown	52,093	77.12	12.71	16.61	24.59	17.48	28.61
No	6,285	68.21	4.96	7.02	15.26	15.56	57.2
Yes	41	61.51	0	4.88	7.32	7.32	80.49
Courses taken							
3011	56,378	76.21	11.93	15.59	23.63	17.23	31.63
Other	2,041	74.36	10.14	15.19	22	18.28	34.39
Gifted/talented							
Unknown	986	67.26	4.97	5.78	15.31	15.72	58.22
No	47,768	73.74	6.49	12.74	24.1	19.64	37.03
Academic	9,054	89.19	39.82	30.85	21.65	5.3	2.39
Artistic	351	80.92	15.38	20.51	31.91	17.38	14.81
Both	260	91.47	47.31	35.38	13.85	2.31	1.15
504 plan							
Unknown	17,370	76.67	12.02	16.26	24.43	17.13	30.15
No	40,420	75.94	11.84	15.31	23.18	17.31	32.36
Yes	629	75.12	9.22	13.51	24.96	18.12	34.18
Alternative school							
Unknown	40,016	76.96	12.53	16.4	24.39	17.57	29.12
No	17,078	75.2	11.1	14.62	22.65	16.83	34.81
Yes	1,325	63.93	1.66	3.09	10.87	13.58	70.79
Accommodations							
No	55,912	76.79	12.37	16.18	24.28	17.51	29.66
Yes	2,507	61.95	0.68	2.03	7.78	11.73	77.78

TABLE 5.4
English 1 Operational Test, Adult Education Programs:
Percentages of Student Scores in Letter-Grade Equivalents, Overall and by Demographics

Demographics	N	Mean Scale Score	A	B	C	D	F
Overall	14	66.64	7.14	0	14.29	14.29	64.29
Gender							
Female	5	71	20	0	0	20	60
Male	9	64.22	0	0	22.22	11.11	66.67
Ethnicity							
African American	7	59.29	0	0	0	14.29	85.71
Hispanic	1	100	100	0	0	0	0
White	5	70.2	0	0	40	20	40
Other	1	67	0	0	0	0	100
Language							
Unknown	4	67.75	0	0	0	50	50
English Speaker I	1	46	0	0	0	0	100
English Speaker II	9	68.44	11.11	0	22.22	0	66.67
Lunch							
Unknown	4	67.75	0	0	0	50	50
No Meals	4	77.5	25	0	50	0	25
Free Meals	6	58.67	0	0	0	0	100
IEP							
Yes	1	67	0	0	0	0	100
No	13	66.62	7.69	0	15.38	15.38	61.54
Migrant							
Unknown	7	71.29	14.29	0	28.57	0	57.14
No	7	62	0	0	0	28.57	71.43
Courses taken							
3011	10	66.2	10	0	20	0	70
Other	4	67.75	0	0	0	50	50
Gifted/talented							
Unknown	4	67.75	0	0	0	50	50
No	10	66.2	10	0	20	0	70
504 plan							
Unknown	10	70.6	10	0	20	20	50
No	4	56.75	0	0	0	0	100
Alternative school							
Unknown	11	70.27	9.09	0	18.18	18.18	54.55
Yes	3	53.33	0	0	0	0	100
Accommodations							
No	13	66.62	7.69	0	15.38	15.38	61.54
Yes	1	67	0	0	0	0	100

TABLE 5.5
Physical Science Operational Test, Regular Schools:
Percentages of Student Scores in Letter-Grade Equivalents, Overall and by Demographics

Demographics	N	Mean Scale Score	A	B	C	D	F
Overall	50,229	72.07	9.28	10.59	16.85	18.19	45.1
Gender							
Unknown	130	65.4	6.15	3.85	7.69	17.69	64.62
Female	24,930	71.76	7.51	10.33	17.41	19.49	45.26
Male	25,169	72.41	11.05	10.88	16.35	16.9	44.83
Grade							
8	418	80.68	13.16	23.21	30.62	17.94	15.07
9	36,932	73.01	10.47	11.51	17.64	18.07	42.3
10	11,431	69.53	6.1	7.76	14.71	18.77	52.66
11	855	66.01	3.04	6.55	9.12	16.02	65.26
12	313	66.43	2.56	4.15	12.14	19.81	61.34
Other	280	62.92	2.14	5.36	8.21	14.29	70
Ethnicity							
Unknown	135	65.19	5.93	2.96	8.15	15.56	67.41
African American	18,844	65.91	2	4.62	11.46	17.85	64.07
American Indian	115	70.92	11.3	8.7	11.3	14.78	53.91
Asian	518	82.34	28.19	19.11	18.53	14.67	19.5
Hawaiin/Pacific Islander	50	77.8	12	18	30	14	26
Hispanic	2,031	68.49	5.27	6.7	14.87	17.04	56.13
White	27,758	76.36	14.18	14.81	20.61	18.66	31.74
Other	778	71.43	8.74	9.9	19.02	15.94	46.4
Language							
Unknown	1,621	61.75	3.21	2.65	6.05	11.35	76.74
Pre-functional	216	57.03	0	0	2.31	6.48	91.2
Beginner	207	59.61	1.93	1.93	3.86	9.18	83.09
Intermediate	257	65.1	3.11	3.89	10.12	13.62	69.26
Advanced	352	71.39	5.11	10.51	15.91	20.45	48.01
Full English Proficient	101	72.22	11.88	7.92	15.84	19.8	44.55
Title III First Year Exited	1	70	0	0	0	100	0
Title III Second + Year Exited	203	73.59	6.4	12.32	18.23	25.62	37.44
English Speaker I	168	73.01	14.29	9.52	14.88	16.67	44.64

TABLE 5.5
Physical Science Operational Test, Regular Schools:
Percentages of Student Scores in Letter-Grade Equivalents, Overall and by Demographics

English Speaker II	47,028	72.58	9.62	10.99	17.4	18.49	43.51
Parent Waiver	75	71.39	6.67	9.33	16	21.33	46.67
Lunch							
Unknown	1,200	69.15	9.75	8.75	11.42	16.42	53.67
No Meals	27,011	76.48	14.45	14.7	20.73	18.63	31.49
Free Meals	18,418	66.15	2.42	5.12	11.55	17.42	63.48
Reduced Meals	3,600	70.2	5.36	8.31	16.69	19.36	50.28
IEP							
Yes	4,240	60.77	1.6	2.08	5.9	10.38	80.05
No	45,989	73.11	9.98	11.37	17.86	18.91	41.88
Migrant							
Unknown	44,801	73.14	10.1	11.4	17.95	18.77	41.78
No	5,400	63.21	2.46	3.89	7.78	13.41	72.46
Yes	28	62.14	0	3.57	14.29	7.14	75
Courses taken							
3211	48,271	72.15	9.29	10.59	16.97	18.32	44.83
3231	27	87.52	25.93	29.63	33.33	11.11	0
3241	150	89.06	38.67	26	28.67	3.33	3.33
Other	1,781	68.26	6.23	8.87	12.35	15.95	56.6
Gifted/talented							
Unknown	613	61.81	2.94	2.45	7.18	12.23	75.2
Academic	43,172	70.26	6.54	8.43	15.84	19.08	50.11
Artistic	5,971	85.61	29.14	26.23	25.05	12.28	7.3
Both	343	76.65	11.95	17.78	17.78	21.28	31.2
No	130	85.27	29.23	29.23	18.46	12.31	10.77
504 plan							
Unknown	14,100	72.59	9.38	11.09	17.33	18.99	43.21
No	35,601	71.87	9.26	10.4	16.66	17.86	45.81
Yes	528	71.64	7.2	9.66	17.42	18.56	47.16
Alternative school							
Unknown	34,157	72.85	9.68	11.21	17.64	18.95	42.51
No	15,140	71.05	8.88	9.74	15.86	17.05	48.47
Yes	932	59.75	0.86	1.5	3.97	8.58	85.09
Accommodations							
No	47,994	72.69	9.66	11.03	17.45	18.65	43.21
Yes	2,235	58.59	1.12	1.16	3.94	8.14	85.64

TABLE 5.6
Physical Science Operational Test, Adult Education Programs: Percentages of Student Scores in Letter-Grade Equivalents, Overall and by Demographics

Demographics	N	Mean Scale Score	A	B	C	D	F
Overall	30	62.23	0	6.67	3.33	6.67	83.33
Gender							
Unknown	1	58	0	0	0	0	100
Female	15	61.47	0	0	0	13.33	86.67
Male	14	63.36	0	14.29	7.14	0	78.57
Ethnicity							
African American	21	59.05	0	0	0	4.76	95.24
White	8	67.75	0	12.5	12.5	12.5	62.5
Other	1	85	0	100	0	0	0
Language							
Unknown	14	59.79	0	7.14	0	7.14	85.71
Initially English Proficient	3	57.67	0	0	0	0	100
English Speaker I	3	67.33	0	0	0	0	100
English Speaker II	10	65.5	0	10	10	10	70
Lunch							
Unknown	10	61	0	10	0	0	90
No Meals	15	64.47	0	6.67	6.67	13.33	73.33
Free Meals	5	58	0	0	0	0	100
IEP							
No	30	62.23	0	6.67	3.33	6.67	83.33
Migrant							
Unknown	26	62.04	0	3.85	3.85	3.85	88.46
No	4	63.5	0	25	0	25	50
Courses taken							
3211	19	64	0	5.26	5.26	10.53	78.95
Other	11	59.18	0	9.09	0	0	90.91
Gifted/talented							
Unknown	9	58.33	0	0	0	0	100
No	16	64.44	0	12.5	6.25	12.5	68.75
Academic	5	62.2	0	0	0	0	100
504 plan							
Unknown	25	63.4	0	8	4	4	84
No	4	56	0	0	0	25	75
Yes	1	58	0	0	0	0	100
Alternative school							
Unknown	27	62.89	0	7.41	3.7	3.7	85.19
No	3	56.33	0	0	0	33.33	66.67
Accommodations							
No	30	62.23	0	6.67	3.33	6.67	83.33

CHAPTER 6

DESCRIPTIVE STATISTICS

Descriptive statistics of scale score distributions for the three test administrations of the current year combined are presented in table 6 for students overall and by gender and race.

TABLE 6
2007-08 EOCEP Test Administration Summary Statistics: Regular Schools and Adult Education Programs, Overall and by Gender, Race, and Accommodations

Algebra 1/Math Tech 2

Regular Schools				Adult Education Programs			
		Scale Score				Scale Score	
	N	Mean	SD		N	Mean	SD
Overall	60,081	78.99	11.8	**Overall**	44	69.45	8.46
Gender				**Gender**			
Female	29,948	78.98	11.59	Female	21	69.48	9.64
Male	29,879	79.08	11.96	Male	23	69.43	7.45
Race				**Race**			
African	22,942	74.15	10.43	African	25	66.6	7.05
White	32,621	82.37	11.43	White	17	73.88	8.77
Accommodations				**Accommodations**			
No	57,804	79.37	11.73	No	40	69.55	8.87
Yes	2,277	69.43	9.26	Yes	4	68.5	1.73

English 1

Regular Schools				Adult Education Programs			
		Scale Score				Scale Score	
	N	Mean	SD		N	Mean	SD
Overall	58,419	76.15	13.49	**Overall**	14	66.64	14.81
Gender				**Gender**			
Female	28,724	77.41	12.88	Female	5	71	17.49
Male	29,457	75	13.92	Male	9	64.22	13.61
Race				**Race**			
African	22,661	70.56	12.03	African	7	59.29	7.87
White	31,407	80.48	12.87	White	5	70.2	15.22
Accommodations				**Accommodations**			
No	55,912	76.79	13.23	No	13	66.62	15.42
Yes	2,507	61.95	11.35	Yes	1	67	

Physical Science

Regular Schools				Adult Education Programs			
		Scale Score				Scale Score	
	N	Mean	SD		N	Mean	SD
Overall	50,229	72.07	14.03	**Overall**	30	62.23	9.46
Gender				**Gender**			
Female	24,930	71.76	13.28	Female	15	61.47	6.23
Male	25,169	72.41	14.72	Male	14	63.36	12.42
Race				**Race**			
African	18,844	65.91	11.72	African	21	59.05	6.37
White	27,758	76.36	13.82	White	8	67.75	10.74
Accommodations				**Accommodations**			
No	47,994	72.69	13.85	No	30	62.23	9.46
Yes	2,235	58.59	10.84	Yes			

CHAPTER 7

RELIABILITY

In this chapter, multiple types of reliability indexes are presented. For the total tests, two measures of the reliability of raw scores and the classical standard error of measurement (SEM) are given. At the passing cut scores, conditional standard errors of measurement (CSEM) for raw scores, for scale scores, and measures of decision consistency were determined.

7.1 RELIABILITY OF RAW SCORES

Table 7.1 reports the reliability coefficients and SEMs. The reliabilities of the total raw scores were computed using the Kuder-Richardson formulas 20 (KR20) and 21 (KR21). The KR21 reliability coefficients were used in computing the CSEM for the raw scores shown below, in section 7.2.

Table 7.1
Reliability Coefficients of Raw Scores

Administration	Number of Items	Number of Test Takers	KR-20	KR-21	Classical SEM
Algebra 1/Mathematics for the Technologies 2					
Fall 2007	50	8,400	0.82	0.80	3.19
Spring 2008	50	49,375	0.88	0.86	3.08
Summer 2008	50	619	0.81	0.80	3.34
English 1					
Fall 2007	55	7,633	0.88	0.88	3.29
Spring 2008	52	48,347	0.89	0.89	3.09
Summer 2008	55	411	0.87	0.85	3.34
Physical Science					
Fall 2007	55	10,811	0.90	0.90	3.31
Spring 2008	55	37,504	0.87	0.86	3.35
Summer 2008	55	148	0.84	0.84	3.49

7.2 OVERALL AND CONDITIONAL SEM

The overall classical SEM is defined as $s_x\sqrt{1-r_{xx}}$, where s_x is the standard deviation of the scale score and r_{xx} is the reliability coefficient. The CSEM for raw scores at the cut score was computed using the following formula (Feldt and Qualls 1998; Huynh, Meyer, and Barton 2000):

$$\text{raw score } CSEM = \sqrt{\left(\frac{1-KR20}{1-KR21}\right)\left(\frac{c(k-c)}{k-1}\right)}, \text{ where } c = \text{cut score and } k = \text{number of items.}$$

The scale score CSEM at the passing cut score was computed on the basis of the conditional standard error of the Rasch ability cut score. The scale score CSEM is defined as the reciprocal of the square root of the test information function at the point on the ability continuum that corresponds to the scale score cut (Hambleton, Swaminathan, and Rogers 1991). Although classical and conditional SEMs serve similar roles, the values of the conditional standard errors are determined separately for each possible test score, while the classical SEM is a single value used for all scores. Table 7.2 presents both the raw score and scale score CSEMs.

TABLE 7.2

2007–08 EOCEP Conditional Standard Errors of Measurement

Administration	Raw Scores	Scale Scores
Algebra/ Mathematics for the Technologies 2		
Fall 2007	3.36	4.21
Spring 2008	3.24	4.14
Summer 2008	3.41	4.04
English 1		
Fall 2007	3.64	3.75
Spring 2008	3.57	3.76
Summer 2008	3.89	3.75
Physical Science		
Fall 2007	3.74	4.67
Spring 2008	3.61	4.71
Summer 2008	3.74	4.76

7.3 CONSISTENCY OF PASSING CUT SCORES

When student performance is reported in a pass or fail category, a reliability index is computed in terms of the probabilities of consistent classification of students, as specified in standard 2.15 in *Standards for Educational and Psychological Testing* (AERA, APA, and NCME 1999). This index takes into consideration the consistency of classifications for the percentage of examinees who would be classified in the same way on a second (hypothetical) EOCEP administration using either the same form or an alternate equivalent form.

A number of procedures are available for estimating classification errors (Livingston and Lewis 1995; Hanson and Brennan 1990; Huynh 1976; Subkoviak 1976).This report uses the *beta* binomial distribution method (Huynh 1979; Huynh, Meyer, and Barton 2000). Table 7.3 presents a summary of agreements between the operational test classifications—that is, the percentages of students who would be consistently classified in the same category (pass or fail) on two equivalent administrations of the test. The consistency index for the passing score is computed for each administration.

TABLE 7.3

2007–08 EOCEP Consistency Index for Passing Scores

Administration	Consistency Index
Algebra/ Mathematics for the Technologies 2	
Fall 2007	0.825
Spring 2008	0.889
Summer 2008	0.832
English 1	
Fall 2007	0.854
Spring 2008	0.877
Summer 2008	0.831
Physical Science	
Fall 2007	0.862
Spring 2008	0.839
Summer 2008	0.823

CHAPTER 8

VALIDITY

Three types of validity evidence are reported for the algebra test forms: test content, item fairness, and internal structure. Evidence of content validity is presented in the item content distribution across domains and the alignment of the current year's EOCEP test items with the state content standards. Evidence of item fairness is examined with the information on differential item functioning (DIF). Evidence of internal structure is provided in correlations among content domains.

8.1 ITEM DISTRIBUTION ACROSS CONTENT DOMAINS

The EOCEP operational and implementation test forms were constructed according to the test specifications and the test blueprints. These items measured the specific assessment standards that were approved by the SCDE. All items in the test forms were reviewed by the content review committee and the sensitivity review committee and were approved by the SCDE. The current year's EOCEP test form specifications are presented in tables 8.1 through 8.3 by subject.

Table 8.1
Item Distribution by Content Domain for Algebra 1/Math Tech 2

Content Domain	Fall	Summer	Spring
EA-1	8	4	7
EA-2	6	8	9
EA-3	9	5	6
EA-4	11	9	10
EA-5	10	20	12
EA-6	5	4	4
NA	1		2
Totals	**50**	**50**	**50**

Table 8.2
Item Distribution by Content Domain for English 1

Strand/Topic	Fall	Summer	Spring
1	27	18	27
2	3	7	3
3	4	6	4
4	10	10	10
6	11	11	11
Totals	**55**	**52**	**55**

TABLE 8.3
Item Distribution by Content Domain for Physical Science

Strand/Topic	Fall	Summer	Spring
1	8	8	8
2	8	7	8
3	8	8	8
4	8	9	8
5	8	8	8
6	8	8	8
7	7	7	7
Totals	**55**	**55**	**55**

8.2 ITEM DEVELOPMENT

All EOCEP items were developed with reference to the South Carolina curriculum standards and measurement guidelines. Various committees reviewed all items; only items approved by these committees and the SCDE were included in the operational forms.

8.3 DIFFERENTIAL ITEM FUNCTIONING

A critical issue in statewide high-stakes testing is whether the test is fair to all test-takers; therefore, an important goal of item and test development is to produce a pool of items that are judged to be free of bias either toward or against any group of students. All EOCEP items were reviewed both for bias and for differential item functioning (DIF).

The sensitivity review committee examined the EOCEP items for potential bias, including language that might disadvantage a particular group, might be considered offensive to members of a particular group, or might present obstacles to a particular group due to factors unrelated to content and processes specified in the standards.

As with other statistical methodologies, there are numerous widely accepted approaches to detecting potential unfairness in test items. Many of these methods fall into the general category of DIF analyses. DIF statistics provide information regarding relative group performance at the item level for gender and ethnic comparisons while controlling for ability. Once an item is flagged for a significant DIF, judgment is used to determine whether the difference in difficulty shown by the DIF index is unfairly related to group membership. The DIF statistics do not necessarily indicate bias or unfairness in an item but may simply show the relative strengths and weaknesses of the two groups being compared after the overall ability that the test is intended to measure has been controlled for.

Procedure:

The procedure that the AIR selected for detecting DIF was the Mantel-Haenszel (MH) chi-square for dichotomous items. The AIR calculated the Mantel-Haenszel statistic (MH D-DIF) for MC items (Holland and Thayer 1988) to measure the degree and magnitude of DIF. The examinee group of interest is the *focal* group, and the group to which performance on the item is being compared is the *reference* group. In this report, the focal groups for DIF were females and African Americans.

Items were separated into one of three categories on the basis of DIF statistics (Holland and Thayer 1988; Dorans and Holland 1993): negligible DIF (category A), intermediate DIF (category B), and large DIF (category C). The items in category C, which exhibit significant DIF, are of primary concern.

Positive values of *delta* indicate that the item is easier for the *focal* group, suggesting that the item favors the *focal* group. A negative value of *delta* indicates that the item is more difficult for the *focal* group. The item classifications are based on the Mantel-Haenszel chi-square and the MH delta (Δ) value as follows:

- The item is classified as C category if the absolute value of the MH delta value (i.e., |Δ|) is significantly greater than 1 and also greater than or equal to 1.5.
- The item is classified as B category if the MH delta value (Δ) is significantly different from 0 and either the absolute value of the MH delta (|Δ|) is less than 1.5 or the absolute value of the MH delta (|Δ)| is not significantly different from 1.
- The item is classified as A category if delta value (Δ) is not significantly different from 0 or the absolute value of delta (|Δ|) is less than or equal to 1.

The data in table 8.5, below, summarize the number of items in DIF categories for the current year's operational test items.

When the operational forms were constructed, all item statistics from the initial field test were reviewed and approved by the SCDE. Due to the large number of items subjected to DIF analyses, erroneous flags could be expected. All flagged items were closely examined by the SCDE. Inclusion of any flagged item on an operational form (i.e., an item classified as C category) was possible only when the SCDE had approved that item.

Table 8.4
Summary of Differential Item Functioning for Operational Items

		Whites/African-Americans			Males/Females		
Administration	**Cat**	**Alg**	**Eng**	**PS**	**Alg**	**Eng**	**PS**
Fall 2007	A+	21	21	23	19	24	22
	A-	27	31	30	30	28	29
	B+	--	2	2	1	1	1
	B-	1	1	--	--	2	2
	C+	1	--	--	--	--	1
	C-	--	--	--	--	--	--
Spring 2008	A+	16	20	20	18	21	23
	A-	30	27	30	29	28	26
	B+	3	1	5	2	2	3
	B-	--	3	--	--	1	3
	C+	1	1	--	1	--	--
	C-	--	--	--	--	--	--
Summer 2008	A+	20	23	26	18	26	28
	A-	22	29	29	29	27	27
	B+	2	--	--	3	--	--
	B-	2	3	--	--	2	--
	C+	3	--	--	--	--	--
	C-	1	--	--	--	--	--

8.4 CORRELATIONS AMONG CONTENT DOMAINS

Evidence of internal structure was examined using correlations among content domains. On the following pages, tables 8.5 through 8.7 report the correlation matrices for the raw scores among content domains for each test.

Table 8.5
Correlations among Domain Scores for
Algebra 1/Math Tech 2

Domain*	EA-1	EA-2	EA-3	EA-4	EA-5	EA-6	NA	Number of Items
				Fall 2007 (N =2828)				
EA-1	1	0.38	0.35	0.43	0.42	0.28	0.09	8
EA-2	—	1	0.42	0.5	0.49	0.32	0.11	6
EA-3	—	—	1	0.48	0.49	0.3	0.13	9
EA-4	—	—	—	1	0.59	0.33	0.11	11
EA-5	—	—	—	—	1	0.32	0.12	10
EA-6	—	—	—	—	—	1	0.12	5
NA	—	—	—	—	—	—	1	1
				Spring 2008 (N = 16,412)				
EA-1	1	0.41	0.34	0.47	0.52	0.33		4
EA-2	—	1	0.41	0.56	0.57	0.43		8
EA-3	—	—	1	0.46	0.5	0.35		5
EA-4	—	—	—	1	0.68	0.45		9
EA-5	—	—	—	—	1	0.48		20
EA-6	—	—	—	—	—	1		4
				Summer 2008 (N =619)				
EA-1	1	0.36	0.32	0.36	0.43	0.2	0.13	7
EA-2	—	1	0.33	0.42	0.41	0.23	0.1	9
EA-3	—	—	1	0.4	0.45	0.22	0.11	6
EA-4	—	—	—	1	0.55	0.27	0.16	10
EA-5	—	—	—	—	1	0.28	0.17	12
EA-6	—	—	—	—	—	1	0.04	4
NA	—	—	—	—	—	—	1	2

* UF = Understanding Functions
LF = Linear Functions
QOF = Quadratic and Other Functions

Table 8.6
Correlations among Domain Scores for English 1

Domain*	1	2	3	4	6	Number of Items
			Fall 2007 (N =2,532)			
1	1	0.49	0.44	0.61	0.63	27
2	—	1	0.27	0.36	0.38	2
3	—	—	1	0.33	0.36	4
4	—	—	—	1	0.47	10
6	—	—	—	—	1	9
			Spring 2008 (N = 5537)			
1	1	0.67	0.62	0.65	0.09	18
2	—	1	0.56	0.59	0.05	8
3	—	—	1	0.51	0.08	6
4	—	—	—	1	0.06	10
6	—	—	—	—	1	10
			Summer 2008 (N = 411)			
1	1	0.41	0.47	0.54	0.41	27
2	—	1	0.24	0.29	0.25	2
3	—	—	1	0.34	0.24	4
4	—	—	—	1	0.39	10
6	—	—	—	—	1	9

* R1 = Reading Comprehension
R2 = Analysis of Texts
R3 = Word Analysis
RS = Research
C1 = Communication
W1 = Writing

Table 8.7
Correlations among Domain Scores for Physical Science

Domain*	1	2	3	4	5	6	7	Number of Items
Fall 2007 (N = 3,639)								
1	1	0.5	0.57	0.51	0.55	0.49	0.49	8
2	—	1	0.59	0.6	0.55	0.51	0.5	8
3	—	—	1	0.61	0.63	0.57	0.56	8
4	—	—	—	1	0.58	0.52	0.52	8
5	—	—	—	—	1	0.56	0.55	8
6	—	—	—	—	—	1	0.49	8
7	—	—	—	—	—	—	1	7
Spring 2008 (N =12,555)								
1	1	0.4	0.39	0.44	0.43	0.44	0.4	8
2	—	1	0.42	0.52	0.49	0.49	0.48	7
3	—	—	1	0.45	0.44	0.44	0.4	8
4	—	—	—	1	0.54	0.57	0.52	9
5	—	—	—	—	1	0.52	0.51	8
6	—	—	—	—	—	1	0.51	8
7	—	—	—	—	—	—	1	7
Summer 2008 (N = 1,485)								
1	1	0.43	0.54	0.39	0.3	0.46	0.37	8
2	—	1	0.46	0.47	0.35	0.37	0.37	8
3	—	—	1	0.52	0.46	0.48	0.53	8
4	—	—	—	1	0.38	0.32	0.4	8
5	—	—	—	—	1	0.3	0.28	8
6	—	—	—	—	—	1	0.46	8
7	—	—	—	—	—	—	1	7

WORKS CITED

AERA, APA, and NCME 1999. *Standards for Educational and Psychological Testing.* Washington, DC: American Educational Research Association.

Dorans, Neil J., and Paul W. Holland. 1993. "DIF Detection and Description: Mantel-Haenszel and Standardization." In *Differential Item Functioning*, edited by Paul W. Holland and Howard Wainer. Hillsdale, NJ: Lawrence Erlbaum Associates.

Feldt, Leonard S., and Audrey L. Qualls. 1998. "Approximating Scale Score Standard Error of Measurement from the Raw Score Standard Error." *Applied Measurement in Education* 11, no. 2:159–77.

Hambleton, Ronald K., Hariharan Swaminathan, and H. Jane Rogers. 1991. *Fundamentals of Item Response Theory.* Newbury Park, CA: Sage Publications.

Hanson, Bradley A., and Robert L. Brennan. 1990. "An Investigation of Classification Consistency Indexes Estimated under Alternative Strong True Score Models." *Journal of Educational Measurement* 27:345–59.

Holland, P. W., and Dorothy T. Thayer. 1988. "Differential Item Performance and the Mantel-Haenszel Procedure." In *Test Validity,* edited by Howard Wainer and Henry I. Braun. Hillsdale, NJ: Lawrence Erlbaum Associates.

Huynh, Huynh. 1976. "On the Reliability of Decisions in Domain-Referenced Testing." *Journal of Educational Measurement* 13:253–64.

———. 1979. "Computational and Statistical Inference for Two Reliability Indices Based on the Beta-Binomial Model." *Journal of Educational Statistics* 4:231–46.

Huynh, H., J. Patrick Meyer III, and Karen Barton. 2000. *Technical Documentation for the 1999 Palmetto Achievement Challenge Tests of English Language Arts and Mathematics, Grades Three through Eight.* Columbia: South Carolina Department of Education.

Linacre, John M., and Benjamin D. Wright. 2003. *WINSTEPS Rasch-Model Computer Program*. Chicago: Mesa Press.

Livingston, Samuel A., and Charles Lewis. 1995. "Estimating the Consistency and Accuracy of Classifications Based on Test Scores." *Journal of Educational Measurement* 32:179–97.

Rasch, Georg. 1960. *Probabilistic Models for Some Intelligence and Attainment Tests.* Copenhagen: Danish Institute for Educational Research.

SCDE. 2007a. *Fall 2007 District Test Coordinator's Supplement* Columbia: South Carolina Department of Education.

———. 2007b. *Test Administration Manual for School Test Coordinators and Test Administrators*. Columbia: South Carolina Department of Education.

———. 2008a. *Spring 2008 District Test Coordinator's Supplement.* Columbia: South Carolina Department of Education.

———. 2008b. *Spring 2008 Test Administration Manual for School Test Coordinators and Test Administrators.* Columbia: South Carolina Department of Education.

———. 2008c. *Summer 2008 District Test Coordinator's Supplement.* Columbia: South Carolina Department of Education.

———. 2008d. *Summer 2008 Test Administration Manual for School Test Coordinators and Test Administrators.* Columbia: South Carolina Department of Education.

Subkoviak, Michael J. 1976. "Estimating Reliability from a Single Administration of a Criterion-Referenced Test." *Journal of Educational Measurement* 13:265–76.

Wright, Benjamin D., and Mark H. Stone. 1979. *Best Test Design*. Chicago: Mesa Press.